Chapter 6 IMAGE PROCESSING - AND GIS

There are two elements of image processing for remote sensing – the techniques used to enhance images for presentation, and those used to extract information. Most work at the 'pixel' level, and many make use of the scene statistics. If the data inhabit more than one spectral dimension (that is, have color), then a broad range of techniques can be applied to reduce the spectral dimension of the data, and extract information. We begin with a prosaic example of a black and white image, chosen to maintain an intuitive grasp of the image at very high magnification.

A Structure of remote sensing data - DN - what is a pixel….

The image of a model shown on the left was chosen for illustration. The small region around the right eye was extracted, and is shown in an expanded view on the right. The data values associated with each 'eye' pixel are given in the table below the figure – the numbers run from 0 to 226, where 0 corresponds to a black pixel, and 226 to a nearly white pixel.

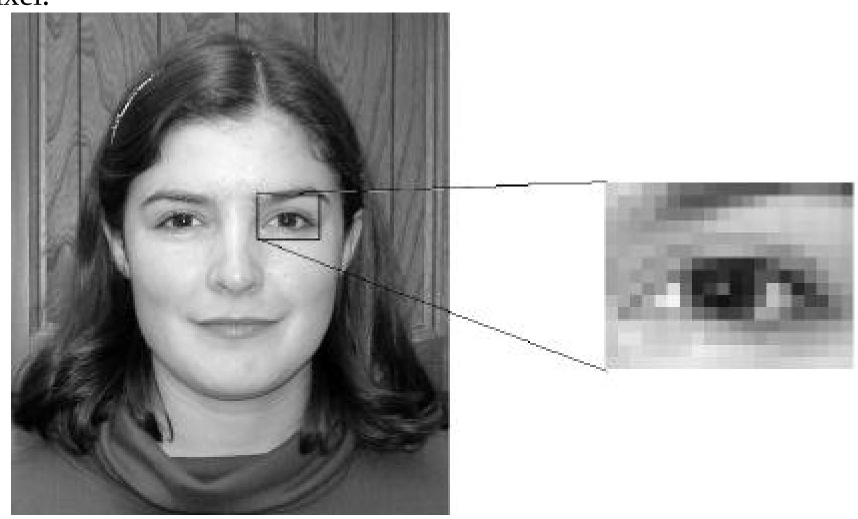


Figure 5-1. Model Susie Olsen

1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	181	188	178	157	153	119	106	107	97	91	91	89	89	87	102	117	119	115	106	82
3	179	160	162	149	132	107	90	86	90	98	114	129	151	172	175	177	169	166	158	141
4	163	158	144	147	120	116	115	121	137	162	174	180	184	184	179	184	182	184	179	170
5	156	149	145	137	139	143	148	156	169	177	179	177	179	182	175	179	177	179	177	169
6	153	151	148	149	153	156	159	152	152	151	153	152	155	162	166	171	173	175	172	166
7	156	152	158	159	150	136	137	146	156	160	158	152	140	134	132	145	161	162	163	158
8	148	158	157	139	144	151	126	87	73	58	55	52	67	96	122	125	123	150	156	153
9	148	152	142	149	143	120	95	48	50	58	43	50	85	85	57	79	111	128	150	152
10	147	152	157	130	143	192	103	47	65	97	38	47	87	165	120	50	71	113	133	144
11	164	153	126	157	197	210	121	71	43	34	44	56	109	170	143	98	73	76	117	132
12	172	134	147	155	151	161	143	110	95	67	71	85	149	146	114	89	99	96	109	131
13	182	187	186	181	175	179	173	171	161	151	134	122	120	116	125	126	129	138	144	153
14	178	198	198	182	179	182	181	191	172	167	162	153	145	153	152	150	152	157	164	169
15	175	185	192	188	185	187	193	205	201	194	190	185	177	173	166	164	170	173	180	182
16	183	185	193	195	198	199	201	200	196	191	188	186	180	180	182	184	187	191	192	189

Many image processing techniques begin with an examination of the statistics of the scene or image, in particular, an important technique is to examine the histogram of occurrences in a scene

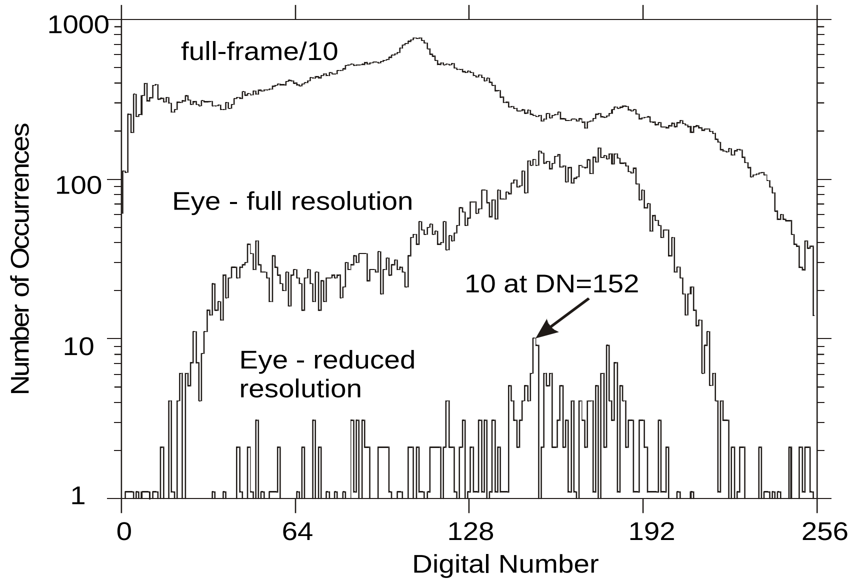


Figure 5-2 Histogram of DN occurrences. The reduced resolution curve corresponds to the image at the right of Figure 5-1, and the table above. The full resolution statistics for the same region are given, along with those for the complete 847×972 image.

This illustration is a little artificial – remote sensing scenes generally have a more even distribution of data values. The IRS-1C image of San Diego illustrated in Chapter 1, for example, has a nice Gaussian shape. Note the large number of near zero values for water.

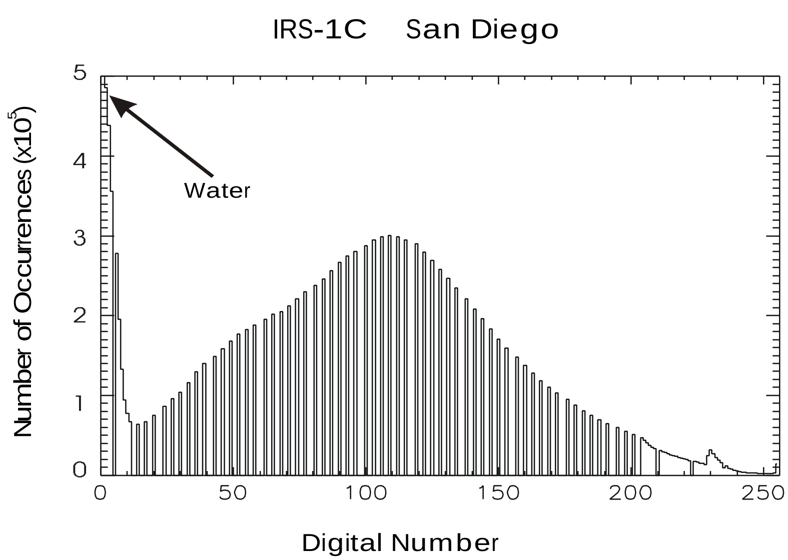


Figure 5-3 Histogram of DN occurrences

B What happens when you have spectral data?

A fairly prosaic example is chosen from a picture of a red rose, taken with an electronic, digital camera. The image is composed of three bands, and can primarily be separated into red and green bands in this scene (plus a specular component reflecting from the leaves). The pixels from the flower lie along the horizontal axis, since the red rose has almost no green to its color. The diagonal line is the ensemble of pixels from the green leaves, starting at zero for the shadowed leaves, and rising up to a DN of 250. A few other spectral components in the scene reflect brown leaves (hidden in the shadows), and a less healthy rose near the top.

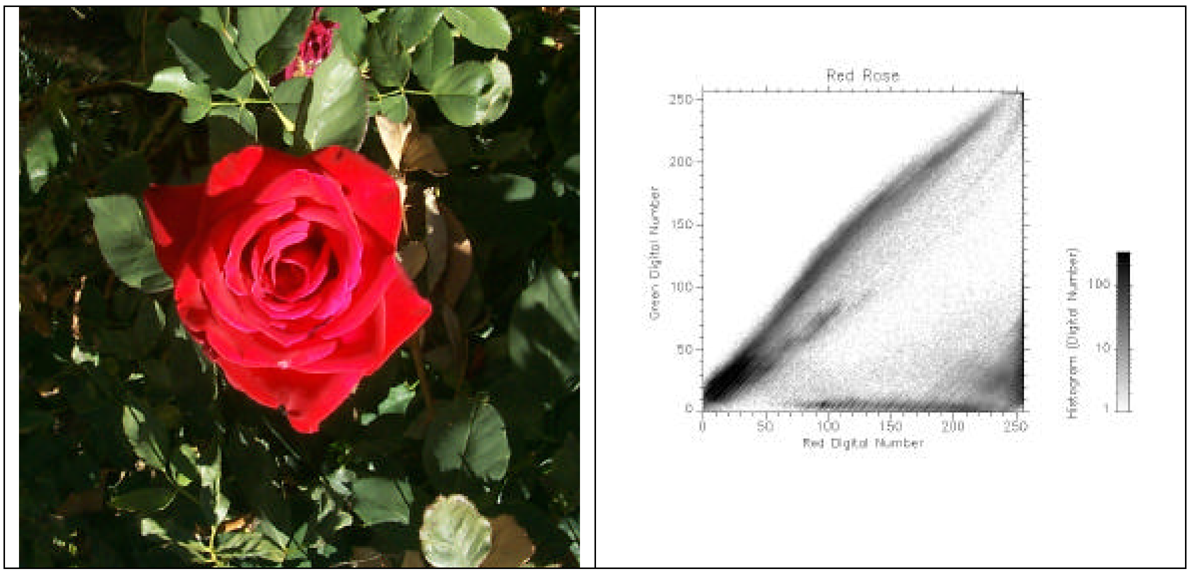


Figure 5-4 (a) Red rose, dark green leaves. (b) Scatter plot of green DN vs. red DN.

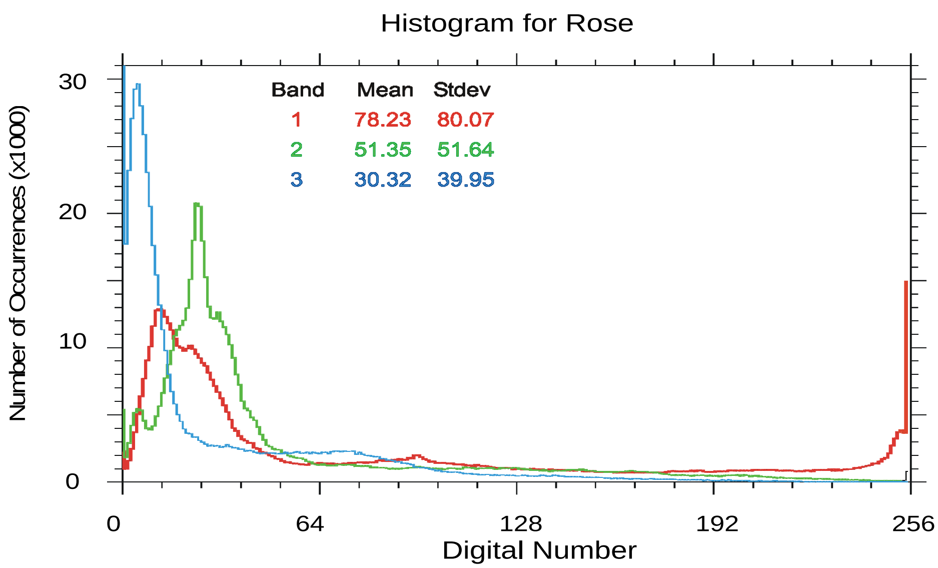


Figure 5-5 Histogram of red (1), green (2), and blue (3) bands from color image above.

One problem with any such data is the highly linear relationship between the different bands. Not really evident here, it becomes much more so when you have higher spectral dimensions to your data (6 or 7 bands with Landsat, 224 bands with AVIRIS). This can partly be understood by using the power of statistical measures such as covariance. If we calculate the covariance matrix for the 583696 points in our rose image, we get:

Covariance Matrix

Band	1	2	3
1	6411.7	1562.4	2127.8
2	1562.4	2666.6	1820.3
3	2127.8	1820.3	1596.0

This is a little hard to interpret without some sort of reference values for what represents large or small variance. The correlation coefficient is self-normalized, and allows us to obtain a measure of the relationship between the different bands.

Correlation Matrix

Band	1	2	3
1	1.000	0.378	0.665
2	0.378	1.000	0.882
3	0.665	0.882	1.000

We can see that the red and green bands are not well correlated (0.378), but green and blue are fairly highly correlated in this image (0.882). It is often helpful to completely decorrelate the different bands by taking just the right combination of bands – what I term a rotation in color space. The rotation that does this is the rotation which diagonalizes the covariance (or correlation) matrix. (Either can be done, which depends upon the nature of the problem at hand). This is called a principal components transform, also known as a Hotelling transform or a Karhunen-Loeve transform. If this is done for the above covariance matrix, the rotation we obtain is:

Eigenvectors

Eigenvector	PC 1	PC 2	PC 3
1 (Red)	0.845	-0.510	0.161
2 (Green)	0.374	0.778	0.505
3 (Blue)	0.383	0.366	-0.848

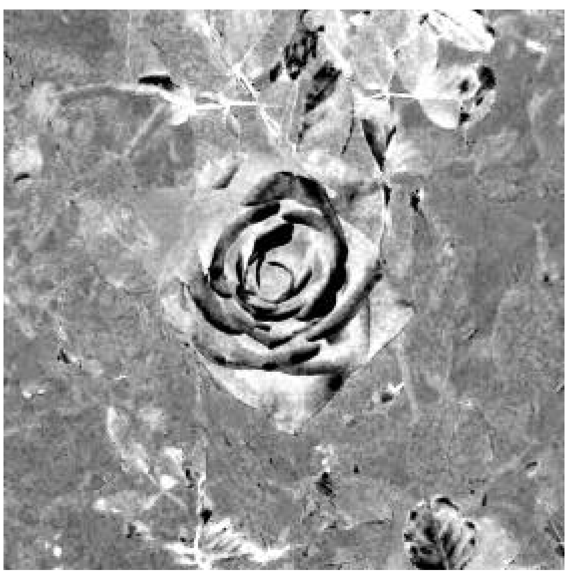
These eigenvectors have the eigenvalues:

PC Band	Eigenvalue
1	8066.72
2	2499.65
3	107.89

The eigenvalues tell us that the variance in the first two principal components is quite large (a large dynamic range in the data values), while the third component has an order of magnitude less variance. The third component is mostly noise. This becomes increasingly true with higher spectral dimensionality in the data, and can be used as a filter for noise in spectral imagers.

Principal Component 1

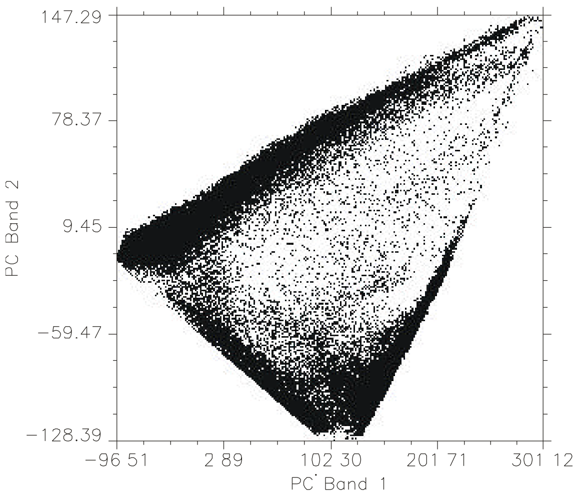
Principal Component 2

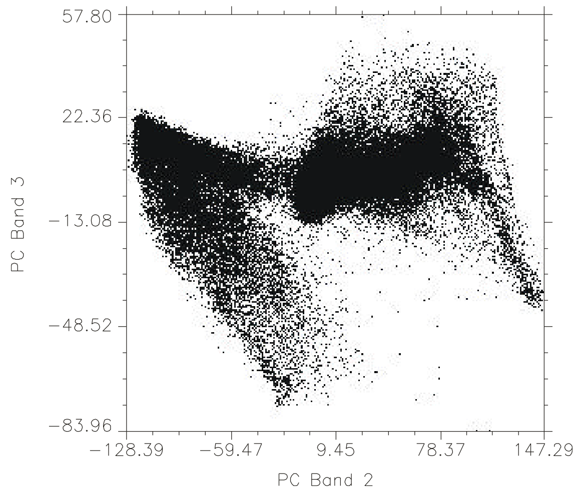


Principal Component 3

The first principal component is nominally the average of the three bands – not unlike a black & white photograph of the scene. For this particular scene, the first component is:
0.845* red + 0.374 * green + 0.383 * blue.
The second principal component breaks out the dark green leaves from the bright red rose:
-0.510* red +0.778*green + 0.366* blue;
basically the difference between green and blue, with things that are bright red very dark, things that are green or blue show up as bright. The third component is sort of the difference between green and blue, and has relatively little information content.

Part of the point, again, was to decorrelate the different bands. If we do a scatter plot in the new principal component space, we see that there is little obvious relationship between the different bands.





C Display vs. Analysis

1 Correcting for sensor artifacts (calibration), atmospheric effects

2 Image processing - histograms, contrast, color tricks

3 spectral processing - rotations in color space, classifiers

D Geographic Information Systems - GI S

E Problems

1. How would you scale a data-set that had a histogram distribution like this: